

DESCRIPTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[Para 1] Not Applicable.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

[Para 2] Not Applicable.

BACKGROUND OF THE INVENTION—FIELD OF THE INVENTION

[Para 3] This invention is a device and method for presenting complex acoustic information, such as music, as visual or tactile information. The acoustic information is processed by a human-like auditory transformation simulating the processing of acoustic information by a human auditory system. The transformed signal is then applied to a tactile or visual presentation. The audience perception of the invention is visual through light, color, animation of an image or object, or touch by movement of an object, providing a synchronicity with the perception of the sound.

BACKGROUND OF THE INVENTION—DESCRIPTON OF RELATED ART

[Para 4] Devices that enhance the human experience of listening to music by expanding the senses used during the experience are popular. Live concerts generally feature motion from the movement of the musicians or an orchestra conductor to the gyrations of a rock band the motion provides an enhancement of the listening experience. The popularity of music video on television, and the popularity of dance are further examples of this combining of listening and motion or visual presentation.

[Para 5] Devices for transforming acoustic information into visual or motion output information are known in the art. In the simplest form these devices simply have a built-in musical tune and a corresponding lighting or color presentation. Examples are U.S. Patent 4,265,159 (Liebman et al.), 5,461,188 (Drago et al.), 5,111,113 (Chu) and 6,604,880 (Huang et al.). A more complex variation is devices that respond to the presence or absence of sound. Examples are U.S. Patents 4,216,454 (Terry), 4,358,754 (Young et al.), 5,121,435 (Chen). Even more complex, is an example that

responds to the intensity of the sound field as described in U.S. Patent 4,440,059 (Hunter).

[Para 6] The circuitry for devices with multiple channels of output use varying forms of electronic circuits to capture the acoustical signal, convert it to an electronic signal, and then divide that signal into non-overlapping frequency bands and drive the presentation device by the signal in a desired frequency band or in multiple bands. Examples of such devices providing a multi-channel light signal in response to the music are in U.S. Patents 3,222,574 (Silvestri, Jr.), 4,000,679 (Norman) 4,928,568 (Snaveley), 5,402,702 (Hata) and 5,501,131 (Hata). Another variation is to take two channels of sound, as is found in stereophonic music signals, and compare the two channels to produce a visual presentation, as taught in U.S. Patent 5,896,457 (Tyrrel). All of these devices work by taking a measurable feature of the sound and using it to provide a presentation of the measurable feature.

[Para 7] Human perception of sound waves (also called sounds in this application) is subjective and is not only a physiological question of features of the ear, but also a psychological issue. For example, there are masking effects that determine if a sound is perceived. A normally audible sound can be masked by another sound. A loud sound will mask a soft sound so that the soft sound is inaudible in the presence of the louder sound. If the sounds are close in frequency the soft sound is more easily masked than if they are far apart in frequency. A soft sound emitted soon after the end of a loud sound is masked by the loud sound, and even the soft sound received just before a loud sound can be masked. Sounds also have many different qualities that the human auditory system can perceive such as tempo, rhythms, intensity variation from highs to lows, and rests of silence.

[Para 8] A visual or tactile presentation that is not representative of the perceived sound does not enhance the audio experience. It instead provides a distraction to the audio experience. On the other hand, if the presentation enhances the audio by responding as the audio is perceived, it enhances the audio experience enabling the audience to visually or tactily experience the tempo, rhythms, intensity variation from highs to lows, and silences of the audio, providing a synchronicity that enriches the combined experience more than either experience individually.

[Para 9] In order to provide a presentation which is representative of the perceived sound, it is necessary to model what humans actually hear. The presentation must represent how sounds are received and mapped into thoughts in the brain, rather than a mere representation of a measurable feature of the sound wave. The presentation also must be capable of displaying a wide range of values representing the wide range of perceptions of sound that human hearing is capable of. What is needed is a presentation that overcomes the limitations of the prior art by seemingly displaying responses to sounds as they occur and reflecting the richness of perceptible components of the sounds such as tempo, rhythms, intensity variation from highs to lows, and silences of the audio, providing a synchronicity with these characteristics.

SUMMARY OF THE INVENTION

[Para 10] This invention is a method and system for providing an audience sound and a visual or tactile presentation that expresses a rich interpretation of acoustic sound, perceived simultaneous with that sound. The method provides for receiving an acoustic signal then performing a human-like auditory transformation of the signal such that the signal has multiple channels reflecting such perceptible qualities as tone, notes, intensities, rhythms and harmonics. A time-sequence scaling of the transformed signals is performed to provide consistency of the presentation, and audience presentation of the transformed signal is provided such that it is perceived simultaneous with the perception of the sound.

[Para 11] The system creates an electronic sound signal from sound waves captured via a microphone, processes the signal with an automatic gain control (AGC) circuit, and converts the analog sound signal to a digital signal using an analog to digital (A/D) circuit. This signal is provided to a processor instructed to perform a human-like auditory transformation on the digital signal such that a multi-channel digital signal representative of human perception of the sound is created. The processor is further instructed to perform a time-sequence scaling of each channel of the multi-channel digital signal to maintain consistency of each signal. These signals are provided to a presentation that uses a multi-channel digital to analog (D/A) circuit to convert the signals, and these analog signals drive a visual or tactile presentation control. The control activates the display such that the presentation provides the audience a visual or tactile presentation of the sound representative of the perception of the sound including characteristics such as tempo, rhythms, intensity variation from highs to lows, and silences of the audio. The system performs the sound signal transformation quickly so the visual or tactile presentation is perceived with the perception of the sound, providing a synchronicity with the sound.

[Para 12] The human-like auditory transformation is made using a human hearing model selected for the presentation desired. Commonly used models are critical bands, mel scale, bark scale, equivalent rectangular bandwidth, and just noticeable difference.

[Para 13] The system may also use analog or digital stored sound signals to produce both the sound and the visual or tactile presentation of the sound. In use with music, the system may also develop an estimate of the music beat. This signal is added to one or more of the visual or tactile presentation channels to enhance the presentation. Types of displays used for the presentation may include multiple channels of lights, multiple color lights, an animated display on a computer or television screen, or projection of the animated display, fountains of water, multiple channels of laser lights, multiple spotlights, motion of an object in multiple degrees of freedom, multiple firework devices, a refreshable Braille display, or vibrating surfaces.

[Para 14] The system may be implemented on an Application Specific Integrated Circuit (ASIC) or a general-purpose computer system, or any other type of digital circuitry that can perform the computer-executable instructions described.

Objects and Advantages

[Para 15] One object of this invention is to provide a visual presentation representative of the human perception of sound such that the human may watch the presentation change with the perception of the sound.

[Para 16] A second object of this invention is to provide motion of an object representative of the human perception of sound such that the human may observe visually the object motion change with the perception of the sound, and/or observe tactilely the motion change with the perception of the sound.

Brief Description of the Several Views of the Drawings

[Para 17] A more complete understanding of the present invention can be obtained by considering the detailed description in conjunction with the accompanying drawings, in which:

[Para 18] Figure 1 is a block diagram of the acoustic presentation system showing the features of the device.

[Para 19] Figure 2 is a block diagram of the acoustic presentation system showing an embodiment containing beat detection.

[Para 20] Figure 3A is a block diagram of the signal reception feature of the device using an external sound source.

[Para 21] Figure 3B is a block diagram of the signal reception feature of the device using sound from an analog storage source or a playback device using digital storage but providing an analog output.

[Para 22] Figure 3C is a block diagram of the signal reception feature of the device using sound from a digital storage source

[Para 23] Figure 4 is a block diagram of the human-like auditory transformation feature of the device.

[Para 24] Figure 5 is a block diagram of the presentation feature of the device.

[Para 25] Figure 6 is a schematic diagram of an implementation of the acoustic presentation system on an ASIC using strings of lights as the presentation.

[Para 26] Figure 7 is a schematic diagram of an implementation of the acoustic presentation system on a general-purpose computer.

[Para 27] Figure 8 is a schematic diagram of an implementation of the acoustic presentation system on an ASIC.

Reference Numerals in Drawings

[Para 28] These reference numbers are used in the drawings to refer to areas or features of the invention.

[Para 29]	40	Sound Source
[Para 30]	50	Signal Reception
[Para 31]	52	Microphone
[Para 32]	54	AGC
[Para 33]	56	A/D
[Para 34]	58	Sound Storage Playback
[Para 35]	60	Sound Presentation
[Para 36]	70	Human-like Auditory Transformation
[Para 37]	72	FFT
[Para 38]	74	Human Hearing Model
[Para 39]	80	Beat Detection
[Para 40]	90	Time-Sequence Scaling
[Para 41]	100	Presentation
[Para 42]	102	Multichannel D/A
[Para 43]	104	Presentation Controls
[Para 44]	106	Presentation Display
[Para 45]	120	Application Specific Integrated Circuit (ASIC)
[Para 46]	122	Power Supply
[Para 47]	124	Rectifier
[Para 48]	140	Computer

DETAILED DESCRIPTION OF THE INVENTION

[Para 49] The present invention is an electronic device and a method of providing a visual or tactile presentation of an acoustic presentation, such as music, on a device to be observed, as the acoustic presentation is perceived. Referring to figure 1, The invention performs four functions, signal reception (50) is the receipt by the device of the acoustic presentation, human-like auditory transformation (70) is the changing of

the signal into channels of acoustic frequency band energy vectors that represent the human perception of the acoustic presentation, time-sequence scaling (90) is the scaling of a time interval of the output presentation to the previous time interval to provide consistency of the presentation, and the presentation (100) is the display of multi-channel lights, colors of lights, display animation, or object animation that moves to display the acoustic presentation, and the controls and signal conditioning needed for the presentation display.

[Para 50] The signal reception (50) is a microphone (52) to convert the sounds coming from the sound source (40) to an electronic sound signal as shown in figure 3A. The signal is processed by the automatic gain control, AGC (54), to maintain the amplitude in a range that can be processed by the analog to digital converter, A/D (56). The digital signal is then provided to a digital processor for the human-like auditory transformation (70). The processor may be a general purpose computer, an Application Specific Integrated Circuit (ASIC), or any other type of digital circuitry that can perform the computer-executable instructions described herein.

[Para 51] The human-like auditory transformation (70) is shown in figure 4. The digital sound signal from the signal reception (50) is fed to the Fast Fourier Transform, FFT (72), which provides a Fourier spectrum frequency domain signal of the time-domain sound signal. The resulting frequency vectors are divided into channels weighted by the human hearing model (74).

[Para 52] Human hearing models (74) are based on studies of human acoustic perception and are known to those skilled in the art of computer voice recognition, where they are applied in modeling speech. Humans do not hear all frequencies the same, so the output of the FFT is combined into frequency bands by one of these models in a number of groups equaling the desired number of presentation channels. Any of several models may be used.

[Para 53] One such model is the critical band. Humans can hear frequencies in the range from 20 Hz to 20,000 Hz, however this range can be divided into experimentally derived critical bands that are non-uniform, non-linear, and dependent on the perceived sound. The critical bands are a series of experimentally derived frequency ranges in which two sounds in the same critical band frequency range are difficult to tell apart, in other words are perceived as one sound. Critical band ranges are used to weight the FFT spectrum of the sound and deliver these to the presentation (90). The number of channels desired for the presentation determines the number of groups.

[Para 54] An alternate model is the bark-scale. The bark scale corresponds to the first 24 critical bands of hearing and is often related to frequency (in hertz) by the relationship:

[Para 55]
$$\text{barks} = 13 * \arctan(0.00076 * f) + 3.5 * \arctan((f/7500)^2)$$

[Para 56] The bark scale may also be replaced with an Equivalent Rectangular Bandwidth (ERB) that decreases the band size of the bark scale at lower frequencies,

below 500 Hz. The ERB was developed to account for the temporal analysis performed by the human brain on speech signals. The ERB is for moderate sound levels:

[Para 57] $ERB = 0.108f + 24.7$

[Para 58] Another model is the Just Noticeable Differences (jnd). The jnd provides band sizes based on the perception of changes in sound frequency, or pitch, that are perceived half the time. The jnd in hertz increases with the initial frequency in accordance with Weber's Law:

[Para 59] $df/f = C$

[Para 60] Still another alternate, the mel scale (m), is based on the perceived frequencies, or pitch, judged by listeners to be equal in distance one from another. It is related to frequency in hertz by the relationship:

[Para 61] $m = 1127.01048 \log (1 + f / 700)$

[Para 62] The signal may be further modified to emphasize the beat of the music as shown in figure 4. The beat detection (80) derives an estimate of the beat by summing the values for all the output levels of the FFT (72). This total energy value is scaled by determining the minimum and maximum values for the current and one previous time step. The minimum is subtracted from the maximum to derive the range of this short time period, and the range of desired output levels is divided by this range to provide a beat factor. The beat component is applied to the value of one or more channels of the human hearing model (74) output, depending on the type of presentation. Some presentations, such as a dancing doll, may not require this emphasis and so this feature may not be applied, or even calculated, in those cases.

[Para 63] The output of the human-like auditory transformation (70) is multiple channels of frequency domain energy values each in a range of desired output values. It is desired these values be in a desired range corresponding to the possible display states for the presentation that is used. These values may have been modified by the beat signal detection as previously described. This output, by channel, is stored in a memory for a time interval on the order of 1 second by the time-sequence scaling (90). This stored information, and the current value are used to derive a scale factor used to maintain the output value within the desired range. The range is calculated from the minimum and maximum of the stored and current time intervals. The desired range in output values for the presentation is divided by this calculated range to develop a scale factor that is applied to the current value.

[Para 64] The presentation (100) in figures 1 and 2 is shown in figure 5. The multiple channel output from the time-sequence scaling (90) is converted by the multichannel D/A (102), digital to analog converter, to an analog signal for operating the presentation controls (104). Example presentation displays (106) that are commonly available include, but are not limited to, multiple channels of lights, multiple color lights, an animated display on a computer or television screen, or projection of the

animated display, fountains of water, multiple channels of laser lights, multiple spotlights, motion of an object, such as a doll, in multiple degrees of freedom, multiple firework devices, a refreshable Braille display, vibrating surfaces, or other device providing visual or tactile information. The presentation controls (104) will vary depending on the type of display selected, but are commonly available for these displays. Power controls may be used for light strings and color displays to control brightness, image generation and motion generation circuits or software for video and computer displays, multiple motor controllers, solenoid valves, or igniters, for displays of motion of one or more objects or devices.

[Para 65] One example of the present invention device is shown in figure 6. This is an Application Specific Integrated Circuit (120) or ASIC processor implementing the method of the present invention to provide a presentation display of multiple strings of lights (106). The ASIC and presentation are powered by the power supply (124) from an AC power source. A microphone is (52) incorporated into the device. The microphone sound signal is provided to the signal reception (50). Signal reception provides a digital signal as previously described to the human-like auditory transformation (70). Beat detection (80) is used for this presentation, and the results of both the human-like auditory transformation (70) and the beat detection (80) are a multi-channel signal with a range of up to 128 light intensity levels, maintained by the time-sequence scaling (90) as previously described.

[Para 66] The ASIC output signals are provided in digital form to the D/A (102) for powering the multiple strings of lights through the presentation controls (104), which control the power applied to the presentation. The resultant presentation is four channels of lighting strings responding in brightness to an acoustic presentation in the vicinity of the device. The four channels of lighting strings respond individually to the acoustic presentation, modeling the perception of the acoustic presentation as heard by the audience.

Other Embodiments

[Para 67] The signal detection may be a stored signal, as shown in figures 3B and 3C, derived from an analog or digital electronic sound storage playback (58) device. Examples of such devices are a computer hard drive, a computer floppy disk, a computer flash memory device, a tape, a compact disk, or other storage device. Figure 3B shows the use of an analog sound storage playback device. The signal is processed by the AGC (54) and the A/D (56) as with the microphone described previously. Digital sound storage playback may be input directly to the human-like auditory transformation (70) as shown in figure 3C.

[Para 68] The sound presentation (60) may include a time delay to accommodate some presentation displays (106) that inherently take additional time to be perceived, such as a fireworks display. The signal is processed and provided to the audience

through an audio playback device. The device may be integral with the computer operating the presentation software, or a separate device to provide special effects, such as surround sound, or to accommodate multiple sound sources for large audiences, as with a fireworks display.

[Para 69] A device for any of the visual or tactile presentations with digital sound storage is shown in figure 7. The general purpose computer (140) stores digital music files on the hard disk. This sound storage provides a selected signal to be processed by the computer using the instructions of the present invention previously described. The output of this processing is a multichannel digital signal from the computer to the multichannel D/A (102). The multiple analog signals of each channel then go to the presentation actuation (104) that implements the presentation display (106). The selected signal also is output from the computer to the sound presentation (60). The sound presentation (60) and the visual or tactile display (106) may then be perceived by the audience together.

[Para 70] A device using an ASIC processor for any of the visual or tactile presentations with sound storage is shown in figure 8. A sound storage playback device (58) provides the sound signal to the ASIC (120) and to the sound presentation (60). Processing of the sound signal to produce the visual or tactile presentation display (106) occurs quickly enough so the audience perceives the sound and presentation simultaneously for most presentations. There are some presentations, such as a fireworks display as noted previously, where a time delay for the sound presentation (60) is necessary to provide the perception the presentation display (106) and the sound presentation is simultaneous.